Comparing Data

The Osca Troubled Families Dataset

This analysis is based on information collected through Freedom of Information requests to Local Authorities throughout England and Wales and progress data released by www.gov.uk on a regular basis as part of the Troubled Families (TF) programme.

Headline findings based on the data are given first. More detail on methodology, further analysis, implications and conclusions follow.

For a shorter introduction to the TF dataset and what it might tell us about caseload, there is a blog post here.

Headline Findings

The data collected through this research project gave much interesting insight into the functioning of the TF programme. It showed that there were substantial differences in the intensity of support provided by different Local Authorities, and wide variations between progress figures released by the government and figures cited by Local Authorities in direct FOIA responses. Key findings are summarized below.

1. There was a significant variation based on the local council's political orientation, with Labour councils generally appointing for more staff members to TF programmes than conservative equivalents (figures i and ii in Analysis section below).

   Average staff numbers for Conservative held councils were 33 in 2013-14 and 36 in 2014-15. Labour held councils by contrast had much higher average staffing figures in the programmes, 59 for 2013-14 and 64 for 2014-15. Average client figures were comparable, so this is not, at least in any straightforward way, a case of Conservative councils having a lower level of need. In fact Conservative held councils had slightly higher client figures than their Labour held counterparts in both years (13-14 Con: 806, Lab: 709; 14-15 Con: 653, Lab: 543).

2. The requests asked programmes to share their average levels of sickness absence across TF programmes. Although there was lower rate of return for responses on this question (63%) than most responses in the dataset, the data was nevertheless revealing. Sickness rates were generally very high, with a dataset average across both years of 6 days per person. This is 2 more days than the national average of 4 according to ONS statistics.

The average in turn masks wide variations with some teams having very high sickness rates indeed, 8 TF teams had average sickness rates of 10 days per person or more.

Curiously Labour held councils that typically had higher staffing levels and consequently lower caseload numbers across both years also had higher sickness absence rates. The average for the dataset was 7 days for Labour held councils running TF programmes, and a national average of 4 days for Conservative held councils.

3. As noted above there were very significant variations between the numbers published by LAs in the government data releases and the equivalent numbers released in response to my FOIA requests.

   In a second round of FOIA queries that sought to explore discrepancies, numbers of families 'worked with' differed between data releases and FOIA responses. For the 13-14 data there was not a single Local Authority that provided usable data that gave the same figure for clients in the FOIA response as had been given through the government data release. The average number of families from the data release was 206, but this figure was a much lower 124 as calculated from FOIA responses.

4. Discrepancies on caseload were particularly surprising. Average caseload varied significantly in calculations based on the government data releases compared to responses to questions on caseload in FOIA requests. Caseload as calculated according to the figures released by the government in 2013-14 was 32 and in 2014-15 was 22. Average caseload cited in response to FOIA was the same across both years, and was a much lower 10.

5. Using the information collected through this research, I calculated a 'capacity' for each programme based on a distribution of direct hours of contact with families. There were wide variations between programmes and the mean should be treated with some caution as a result, but it is nevertheless interesting. In 2013-14 the mean number of hours that a TF key worker could spend with a client family was 2 hours. In 2014-15 this had gone up to 3 hours.

6. The dataset also ascribes a 'success' rating to programmes based on the turned around figures as a percentage of the total number of families 'worked with'. These gave an average 'success' rating of 40% for 2013-14 and 57% for 2014-15.

All of these findings, and what they might mean for TF programmes, data transparency and the measurement of social programmes are discussed in more detail below.

Introduction

Over the last few months I've been slowly and painfully building a dataset on the Troubled Families programmes being run across England and Wales.

The underlying reason for doing this bit of building was a niggling feeling that the recent focus on all things innovative has left us with social programmes that we know mainly through their differences, to such a point that we have become unwilling to think about their commonalities.

It sometimes looks rather like the special sauce of the delivery organisation has become much more important than the substantive meat of delivery that it covers, however delectably. It's the sauce that sells. The meat is just so much dead animal, but the sauce can be branded and sold profitably to donors, celebrities and the media. Kids Company, to name a recent example of this approach, had a very solid run of sales of its high sugar special sauce to the point that people stopped asking what kind of meat it was covering, and whether it might possibly be on the turn.

I wanted to get back to the meat, and to see whether social programmes were serving up the same kind of flesh, and whether we could judge its quality sauce-free.

There is an ongoing argument amongst many within the broader social sector about what success might look like and how we can be sure we've seen it in a way that takes us beyond the evidence of conviction. At the risk of caricaturing a more nuanced range of views, this part of it can look – broadly – something like this: commissioners in one corner aiming for standardization of measurement, and the voluntary sector in the other resisting the introduction of comparable data as if their survival depended on it.

This is a complex area, and some suggested approaches certainly look surprising, the SE Ratio and Impact Genome being top of the list of measurement tactics whose ambition oversteps their potential usefulness and reliability and which certainly wildly overshoot their theoretical foundations.

There are, though, good reasons to spend a bit of time and energy on the common traits of social programmes rather than the differences. These need not necessarily lead us to the point of deciding that every single programme creates social value in exactly the same directly comparable way and that this can be usefully captured in a live feed.

The first reason is that, despite rhetorical flourishes around USP, most social programmes just do share common features. This is not a bad thing. It's a basic point about providing support to those who need it. There are only so many ways of doing that. To narrow down already, and get us away from the universal social value creation of SE Ratio type approaches, we might think of relational support and how this can be provided.

Broadly speaking there are four ways in which to provide support through strong relationships:

These are:

1. 1:1 dedicated support
2. Group support
3. Remote support via phone or internet
4. Mixed support offering a combination of these things

Within these basic premises are many different kinds of possible combinations. 1:1 support may be provided by professionals or by volunteers. It may be regular or occasional, intensive or informal. It may lead to specific outcomes such as qualifications, or may be open-ended. Assessing these particular kinds of combinations, however, is more effective if we know something about the basics – for instance how many hours of support are being provided through 1:1 programmes – not least because it is knowing about the basics that will allow us to compare across different providers more fairly and find out – for instance – whether some programmes are more time-intensive than others.

We could – and I'm going to argue, should – do better at sharing measurement as a means of internal appraisal and learning, and of external reporting. This would help organisations to learn from each other, and provide much better baselines for realistic support that would discourage the inflation of programme numbers, a common problem when detailed and testable information is lacking.

Rather than make this a purely theoretical argument, I'm going to offer a fairly thorough analysis of a relational support programme that publishes a certain amount of data online through the gov.uk portal. This is the Troubled Families (TF) programme that has been running for just over two years across England and Wales.

The analysis of the TF dataset outlined here will lead to a further discussion of how benchmarking could be developed as a useful way of assessing organizational capacity, staff workload and delivery quality in a simple and straightforward way. This would not substitute for robust evaluation, but it would support it and buttressed by excellent research I will go on to argue that it could even be developed into a system that had some useful predictive elements.

Background

There are already some initiatives that allow for the rigorous comparison of outcomes using available data already collected by government services. NPC's datalabs are good examples of this new development that show real potential. To really work though, these kinds of datasets need to be matched with further information that can tell us more about the internal structures of the organisations delivering comparable outcomes cross-checked through secondary data sources. Ideally this would go beyond a simple cost-benefit analysis and look at key features of service delivery. By looking at the basic things that programmes are getting right, we could identify when these are consistent features that all programmes could usefully adopt. We would also give funders and commissioners a range of valuable tools for assessing the viability of new programmes, and providers a good, simple and accessible evidence base for designing them.

There might, for instance, be a minimal, baseline level of 1:1 support that all programmes should provide. This level might turn out to vary depending on geography, or demographics or model of delivery, but collecting the data will give us the opportunity to look for exactly those kinds of patterns and see if we can define them in useful ways. The current tendency is to assume that social programmes are best treated like black boxes in which only inputs, outputs and outcomes are measured and shared and in which programme design is considered largely as a creative domain uninformed by evidence. This seems like a missed opportunity. There is evidence already collected that could inform programme design and delivery, and which gives genuinely valuable insight into the basic elements of 1:1 support.

To explore this is more detail I decided to analyse a government funded 1:1 support programme that had been recently rolled out with a particular fanfare and an open commitment to data transparency: the Troubled Families programme. The TF programme publishes regular tables charting success assessed by the number of families 'turned around'. It operates across the whole of England and Wales and therefore offers a relatively rich comparative field. It is also run through local authorities who are subject to the Freedom of Information Act. This means that data can be obtained through a simple – if protracted – accessible process.

The Process

All but one – to Southwark Council – of the requests were made online using the Freedom of Information portal What Do They Know. All requests – except the request to Southwark Council – can be viewed on the public account here.

The requests were made in two series. The first request was sent out to 111 Local Authorities and asked the following questions to all councils (including Southwark):

1. How many Lead Professionals / Keyworkers / Dedicated Workers (ie. 1-2-1 support staff) across council staff and/or commissioned providers are working with the TF families for the years
    a. 2013-2014
    b. 2014-2015?
2. What is the percentage of time spent by Lead Professionals / Keyworkers on work not related to their allocated families (eg. meetings, training, general admin)?
3. What is the average caseload of each Lead Professional / Keyworker (i.e. the number of families with which they are working at any one time)?
4. Do Lead Professionals / Keyworkers regularly work overtime and how many hours of overtime per week on average?
5. What is the average number of sick days taken by Lead Professionals / Keyworkers working on the TF programmes?

The second series was sent to a further 37 Local Authorities who were not contacted in the first round, and asked a slightly different series of questions. These final 37 requests were submitted after the initial round of questions had returned information and contained questions informed by the data from the first round. This was done in order to have a relatively large group of councils who could be contacted with a new set of questions that tested some of the findings and queries that arose from the build of the initial dataset.

The second series of questions was:

1. How many Lead Professionals/Keyworkers/Dedicated Workers (ie.1-2-1 support staff) across council staff and/or commissioned providers worked with identified Troubled Families (TF) for the years
    a. 2013-2014
    b. 2014-2015
2. How many of these Leader Professionals/Keyworkers/Dedicated Workers worked full-time/exclusively with TF families
3. What is the number of families worked with in the years
    a. 2013-2014
    b. 2014-2015
4. What is the average caseload of each Lead Professional/Keyworker/Dedicated Worker working full-time with TF families in the years
    a. 2013-2014
    b. 2014-2015
5. What is the average caseload of each Lead Professional/Keyworker/Dedicated Worker working part-time with TF families for the years

    a. 2013-14
    b. 2014-15
6. What is the number of families entering the TF programme in the years
    a. 2013-2014
    b. 2014-2015
7. What is the number of families exiting the TF programme having been 'turned around' in the years
    a. 2013-2014
    b. 2014-2015
8. What is the percentage of time spent by Lead Professionals/Keyworkers on work not related to their allocated families (eg. meetings, training, general admin)
9. What is the average number of sick days taken by Lead Professionals/Keyworkers working on the TF programmes?

This second series of questions sought to gain more accurate data specifically about the families engaged in each year of the TF programme as well as staff numbers. It was also hoped that this second round of questioning would provide a more revealing insight into the – numerous – discrepancies and surprising findings that came from the analysis of the first round of answers.

Data from the second series of requests is still coming in and only a small sample of 8 councils is included here. As further valid data is received this will be fed into the dataset and updated online.

The Data

There are some big caveats about the data that came back from the Local Authorities. Many of these caveats are the inevitable result of a FOIA research process that passes through various hands and may be interpreted differently by different people. In this case the kinds of problems that affect the data include:

1. Different ways of amalgamating numbers across full-time and part-time Troubled Families staff
2. Differing responses from commissioned providers and in-house providers within LAs complicating aggregation of results
3. Complications due to the three tier approach to classifying families within the TF programme – this introduces specific issues about averaging staff and caseload figures since many LAs differentiate between caseloads for intensive, light and superlight levels of need in accordance with Department for Communities and Local Government (DCLG) guidance
4. Gaps in data collection due to external commissioning processes
5. Withholding of data on privacy grounds (Section 40 refusals)
6. Councils refusing the request altogether, usually claiming that the data requested was not systemically collected already and would require more than £450 to retrieve
7. Councils not collecting the data at all

(A table of Local Authorities who refused to provide the data in this research is given in the appendix, with their reasons for refusal.)

Some of the data returned in the requests point clearly to the imperfect reliability of the TF reporting. This is clear in some of the discrepancies between the data tables published as progress reports on www.gov.uk and the specific figures obtained through the FOIA requests. To take a specific example: the progress report summarizing the TF programme at the end of March 2014 gives a total number of families for each LA, a number of families identified, and a number of families engaged. For Dorset and Redbridge these numbers were as follows:

LA	Total Families	Identified Families	Engaged Families
Dorset	590	479	430
Redbridge	550	550	500

In the FOIA response, however, Dorset claimed only 119 families engaged during 2013/14, whilst Redbridge referred to 256 families. This makes for a significant variation in the dataset calculations, as the table below shows:

LA	Total Staff	Total Families Engaged	Average Caseload
Dorset Progress Data	53	430	8
Dorset FOIA data	53	119	2

Redbridge Progress Data	49	500	10
Redbridge FOIA Data	49	256	5

Both Dorset and Redbridge volunteered these figures in their responses to our FOIA requests. The first round of queries did not directly ask for family numbers in each year. The discrepancies raised by these numbers, however, informed the longer series of questions in the second round of FOIA submissions. In the second round LAs were specifically asked the number of families engaged in both years. This was intended to flag up whether Dorset and Redbridge's figures were anomalies, outliers or errors.

In the small dataset of 14 LAs who have so far returned usable data in the second round of queries and which are cross-referenced here, there *are* additional discrepancies between the data provided in the government data release and figures shared through the FOIA requests.

In every single one of the answers to the FOIA request the figure for the number of clients was different to the number in the government data release. In all cases but one, the data release figure was higher than the FOIA figure, with an average discrepancy of 109 families.

This inconsistency should be enough on its own to give us reason to doubt the figures released by the government concerning the Troubled Families programme. There may be differences of interpretation at work here, or, alternatively, there may be some very problematic inaccuracies and inflation of figures in the government data releases.

The second round of questions sought to explore this point by asking LAs for data on families 'turned around' as part of their TF programmes. By using exactly the same wording as the DCLG, it was hoped that the risk of variation in interpreting the question would be lessened. But here too we see significant variation. This is the data on families 'turned around' for 2013-14 for the group of 14 LAs:

Organisation	**Families 'turned around' from data release**	**Families exiting 'turned around' from FOIA**
AVERAGE	**86**	**83**
Central Bedfordshire	41	57
Darlington	93	130
Bedford	99	103

South Gloucestershire	84	84
Richmond	90	63
Rutland	10	12
Hartlepool	156	97
Solihull	117	117
Harrow	105	not given
City of London	0	1
Wokingham	46	20
Isle of Wight	32	not given
Trafford	181	160
Kingston Upon Thames	52	49

In this case two Local Authorities did return the same figure for families 'turned around' in the data release and the FOIA request. The majority, however, did not.

This long introduction is intended to demonstrate the fundamental problem or reliability that comes with data derived from the Freedom of Information Act. It is hard to determine to what extent the data is trustworthy or accurate. It's even harder to pinpoint the source of any inaccuracy or to determine whether FOIA requests or official data releases are better guides to what is 'really happening' on the ground through the TF programmes. That said, whilst neither is wholly reliable, I'm inclined in this case to give more credence to many of the FOIA figures which tally more closely with the TF guidance supplied by the DCLG, particularly when it comes to caseload.

Right up front though here are four thoughts that stem from this work:

- Data transparency is not enough
- Data quality matters just as much as transparent publication
- Data needs to be tested to determine whether it is accurate and usable
- Well-understood standardisation is essential to the production of useful data that can be analysed effectively and re-used by the sector

Testing the Data

Given resource and time constraints, it wasn't possible to look up every single council's published sources on their TF programmes. What we did do, however, to cross-check our findings, was to randomly select 8 LAs from our dataset and look at these in more detail. In doing this we were specifically looking for the following:

1. Information on their method of delivery – was this mixed in-house and commissioned, or delivered solely through one or the other?
2. Was there other public information that confirmed the numbers provided through the FOIA requests?
3. Was the data that we had received from FOIA bolstered by any further relevant data? eg. any indication of 'churn' figures: new entrants to the programme and returns, further information on the data submitted to the government directly for inclusion in the TF progress information
4. Any other relevant findings

The findings from this cross-check are summarized in this table on the data page of the site. We did not find that this exercise revealed much useful information. Although many LAs do publish guidance on their application of the TF programme, they don't publish cross-referenced figures with the same rigour. There was rather little useful reporting on the websites of the councils we searched, and no significant data.

Analysis

This section will go into the headline findings set out above in more detail, exploring the methodology and the potential conclusions that can be drawn from analyzing the TF dataset.

1. **Political Affiliation**

   As noted in the methodology section above, there are very clear problems with data collected through Freedom of Information requests. These are inevitable when there is no commonly agreed formula for measuring a programme or settled definitions for key aspects of its delivery. This may make for all kinds of interesting innovation, variation and learning, but it also makes comparing programmes tricky. That being the case, it seemed useful to try and explore whether there were any objective influences that might affect the implementation of Troubled Families programmes in different locations and whether those influences would be noticeable in the data.

   There are precedents for using FOIA to assess differential implementation of priority programmes at local level based on the political persuasion of the council. House building has been analysed through this lens, and so have parking costs and council officer pay, amongst others. It was plausible, therefore, to think that this might affect the delivery of TF programmes.

   Sorting the data according to the council's politics did indeed reveal variation in staffing and caseload.

   Labour councils generally appointed far more staff members to TF programmes than conservative equivalents (see figures i and ii below).

   Average staff numbers for Conservative held councils in the dataset were 33 in 2013-14 and 36 in 2014-15. Labour held councils by contrast had much higher average staffing figures in their programmes, 59 for 2013-14 and 64 for 2014-15. Average client figures were comparable, so this is not, at least in any straightforward way, a case of Conservative councils having a lower level of need. In fact Conservative held councils had slightly higher client figures than their Labour held counterparts in both years (13-14 Con: 806, Lab: 709; 14-15 Con: 653, Lab: 543).

   Average staff numbers were lowest overall in councils with No Overall Control (NOC). These councils had averages of 30 and 31 in 2013-14 and 2014-15 respectively.

   Curiously – and this will be discussed in more detail below – there was no significant variation in 'success' rates between Conservative and Labour held councils. In 13-14 Conservative councils had 'success' rates of 38%, and Labour councils of 39%, both Conservative and Labour held councils had success rates of 58% in 2014-15. NOC councils had rates of 42% and 57% for the two years.

The most surprising finding in the data when analysed according to party political affiliation, was the sickness absence rate.

Here there was a clear difference between Conservative and Labour held councils. Labour held councils that typically had higher staffing levels and consequently lower caseload numbers across both years also had significantly higher sickness absence rates. The average for the dataset was 7 days per person for Labour held councils running TF programmes, and the national average of 4 days for Conservative held councils. 60% of Conservative councils who returned sickness absence data had rates under the national average, compared to only 29% of Labour councils who returned this data.

NOC councils had the dataset overall average of 6 days of sickness absence per person. 44% of NOC councils who returned sickness absence data had rates below the national average.

Council	Party Control 13-15	Clients 13-14	Clients 14-15	Staff Numbers 13-14	Staff Numbers 14-15
AVERAGE		**806**	**653**	**33**	**36**
Barnet	CON	705	402	45	45
Bexley	CON	400	262	27	27
Bromley	CON	454	328	4	9
Buckinghamshire	CON	545	366	40	40
Cheshire West and Chester	CON	462	354	12	47
Devon	CON	1016	1138	32	42
Dorset	CON	430	385	53	64
East Riding of Yorkshire	CON	436	382	13.5	19.5
Essex	CON	1531	1723	110	110
Hampshire	CON	1590	1202	30	30
Hertfordshire	CON	1128	848	11	11
Kent	CON	2153	2000	84	84
Leicestershire	CON	810	177	51	51
Medway	CON	430	423	14	13
North Yorkshire	CON	657	490	43	43
Shropshire	CON	437	375	14	10
Staffordshire	CON	1070	986	25	25
Surrey	CON	858	638	32	32
Wandsworth	CON	595	395	10	10.5
West Sussex	CON	998	729	43	47
Westminster	CON	592	405	10	8
Wiltshire	CON	510	297	17	19
Worcestershire	CON	727	705	32	32

Council	Party Control 13-15	Clients 13-14	Clients 14-15	Staff Numbers 13-14	Staff Numbers 14-15
AVERAGE		**709**	**543**	**59**	**64**
Barking and Dagenham	LAB	530	456	46	33
Barnsley	LAB	579	480	20	21
Blackpool	LAB	515	256	9	43
Bolton	LAB	630	580	75	75
Brent	LAB	554	632	51	66
Coventry	LAB	887	689	24	26
Croydon	LAB	622	537	18	18
Derby	LAB	505	428	72	114
Dudley	LAB	720	427	13	15
Gateshead	LAB	498	410	33	41
Haringey	LAB	686	538	7	9
Islington	LAB	767	535	408	408
Knowsley	LAB	485	438	28	28
Lambeth	LAB	975	635	7	17
Leeds	LAB	2190	1306	40	40
Leicester	LAB	943	794	205	144
Newham	LAB	890	844	35	35
North Tyneside	LAB	407	316	24	34.5
Oldham	LAB	586	394	19	18
Plymouth	LAB	650	330	12	14
Rochdale	LAB	587	448	14	14
Rotherham	LAB	644	455	164	187
Salford	LAB	770	542	45	45
Sefton	LAB	650	480	92	116
Sheffield	LAB	1520	1030	49	49
South Tyneside	LAB	368	298	114	114
Southampton	LAB	685	281	34	34
Southwark	LAB	667	748	7	7

St. Helens	LAB	366	415	12	12
Sunderland	LAB	754	567	110	110
Tameside	LAB	568	325	43	43
Waltham Forest	LAB	640	522	40	53
Wirral	LAB	674	621	71	128
Wolverhampton	LAB	582	710	65	70.5

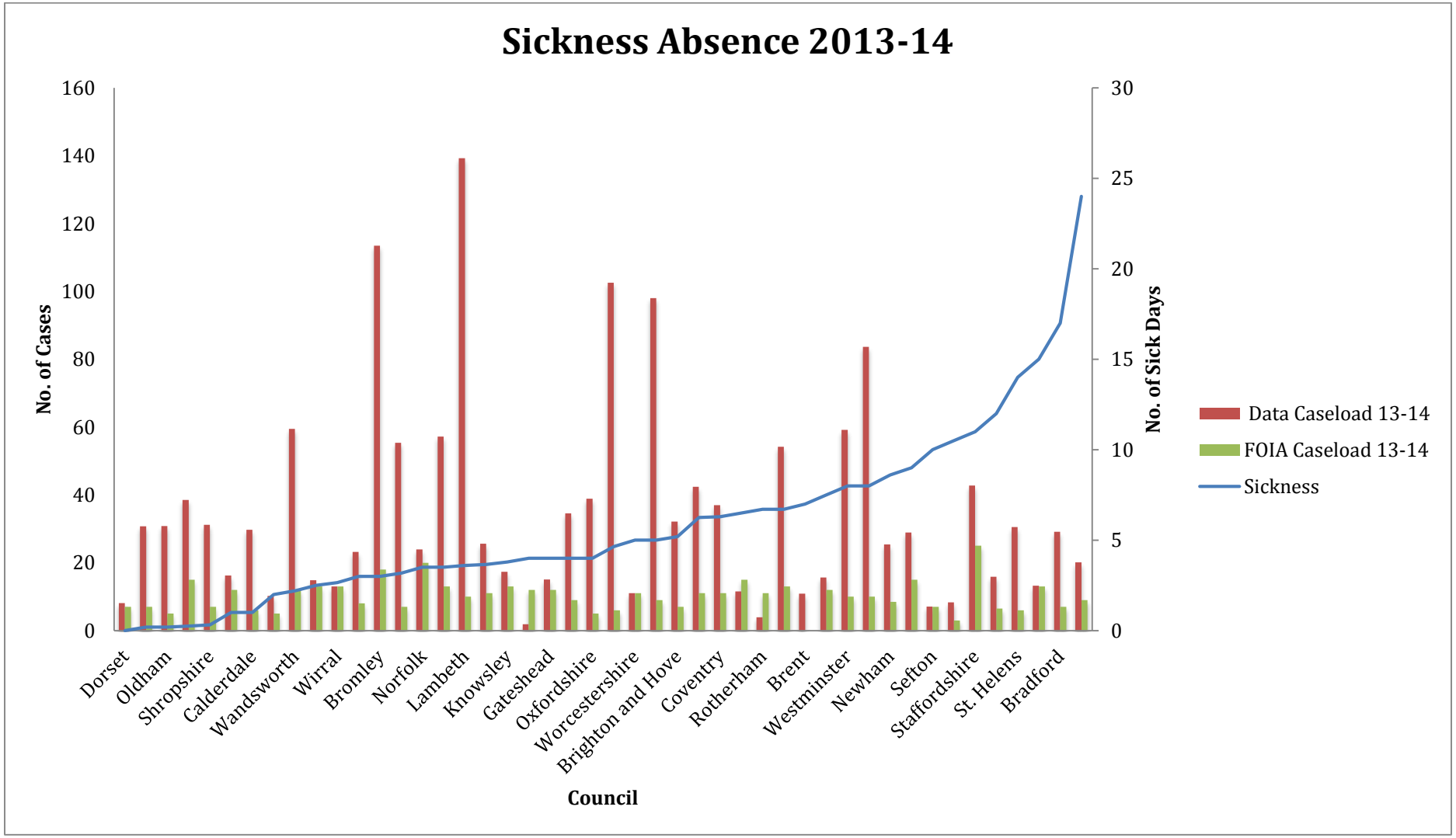
2. Sickness Absence

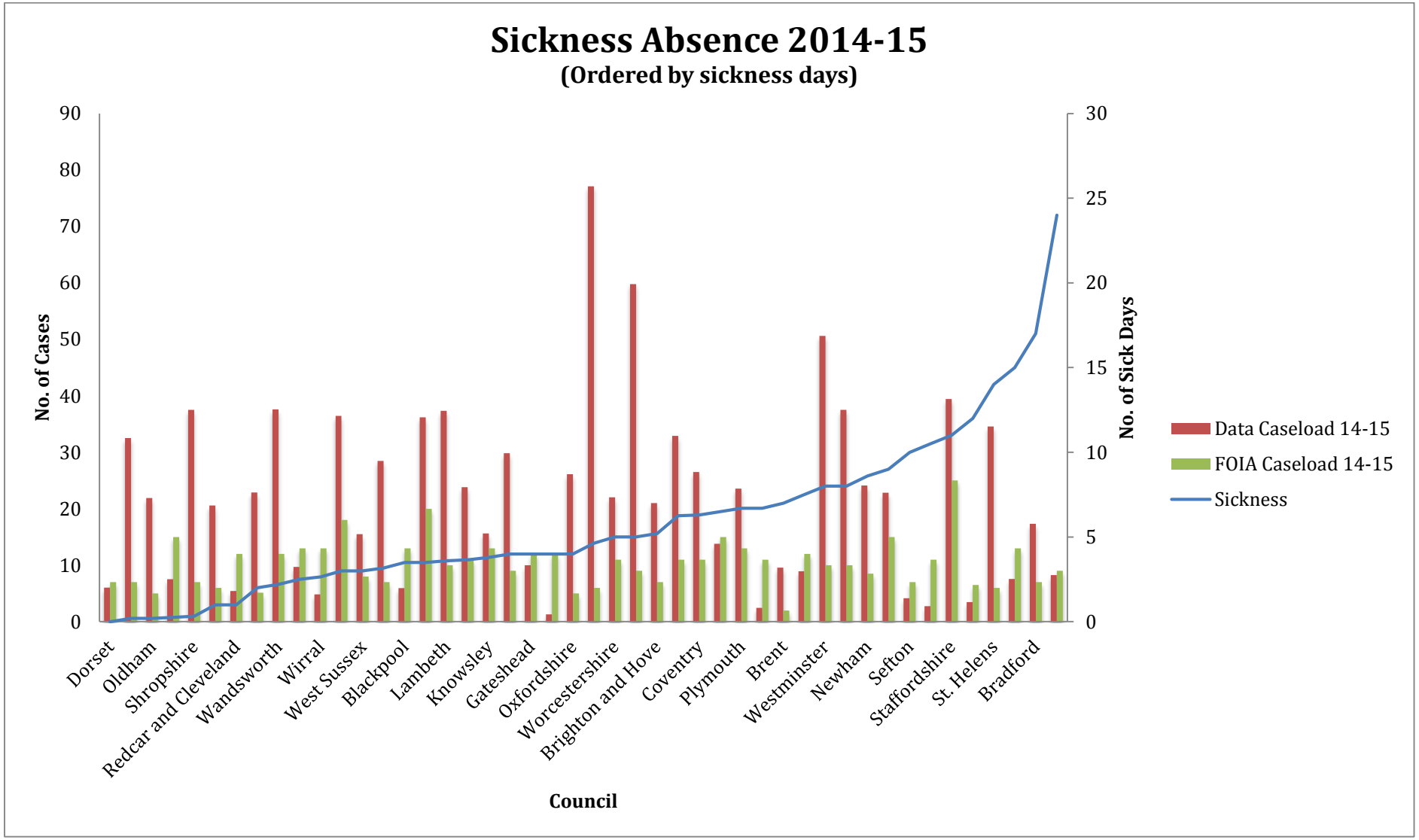
The requests asked programmes to share their average levels of sickness absence across TF programmes. Although there was lower rate of return for responses on this question (63%) than most responses in the dataset, the data was nevertheless revealing. Sickness rates were generally very high across the whole dataset, with a dataset average across both years of 6 days per person. This is 2 more days than the national average of 4 according to ONS statistics.

The average in turn masks wide variations with some teams having very high sickness rates indeed. Eight TF teams had average sickness rates of 10 days per person or more. The highest sickness absence rate in the data set was of 24 days for Southampton. It is of course possible that this figure includes staff members who were long-term sick where generally – wherever possible – long term sickness was excluded from the dataset. But even setting aside potential errors at the higher end of the scale, the data tend to suggest that many TF programmes are operating with depleted staff both in terms of numbers of workers active at the frontline, and their capacity to manage their workloads.

As noted above in discussing the influence of the council's politics on the breakdown of statistics on sickness, rates of absence do not appear to be related in any straightforward way to staffing levels, client numbers or caseload. The charts below show sickness absence rates plotted against caseload figures calculated according to staff numbers and client figures from the government data releases, and the lower caseload figures as given directly in the FOIA responses. In neither case, for either year, is there any obvious correlation.

It has already been suggested that caseload numbers are not sufficient on their own to account for staff experience in social work. If the data collected in the TF dataset do not show a consistent link between sickness and caseload, but do show a link between party leadership of the local council and absence rates, we may need to look more closely at work cultures and, perhaps, levels of control and support to explain high rates of sickness in some TF teams.


Sickness Absence 2013-14
No. of Cases
No. of Sick Days
160
140
120
100
80
60
40
20
0
30
25
20
15
10
5
0
Dorset
Oldham
Shropshire
Calderdale
Wandsworth
Wirral
Bromley
Norfolk
Lambeth
Knowsley
Gateshead
Oxfordshire
Worcestershire
Brighton and Hove
Coventry
Rotherham
Brent
Westminster
Newham
Sefton
Staffordshire
St. Helens
Bradford
Council
Data Caseload 13-14
FOIA Caseload 13-14
Sickness


Sickness Absence 2014-15
(Ordered by sickness days)
No. of Cases
No. of Sick Days
Council
Data Caseload 14-15
FOIA Caseload 14-15
Sickness
0
10
20
30
40
50
60
70
80
90
0
5
10
15
20
25
30
Dorset
Oldham
Shropshire
Redcar and Cleveland
Wandsworth
Wirral
West Sussex
Blackpool
Lambeth
Knowsley
Gateshead
Oxfordshire
Worcestershire
Brighton and Hove
Coventry
Plymouth
Brent
Westminster
Newham
Sefton
Staffordshire
St. Helens
Bradford

3. Caseload

In putting together this dataset a key aim was to test out the idea that caseload numbers and programme capacity might be significant variables in assessing the effectiveness of TF programmes and other social programmes working in similar ways.

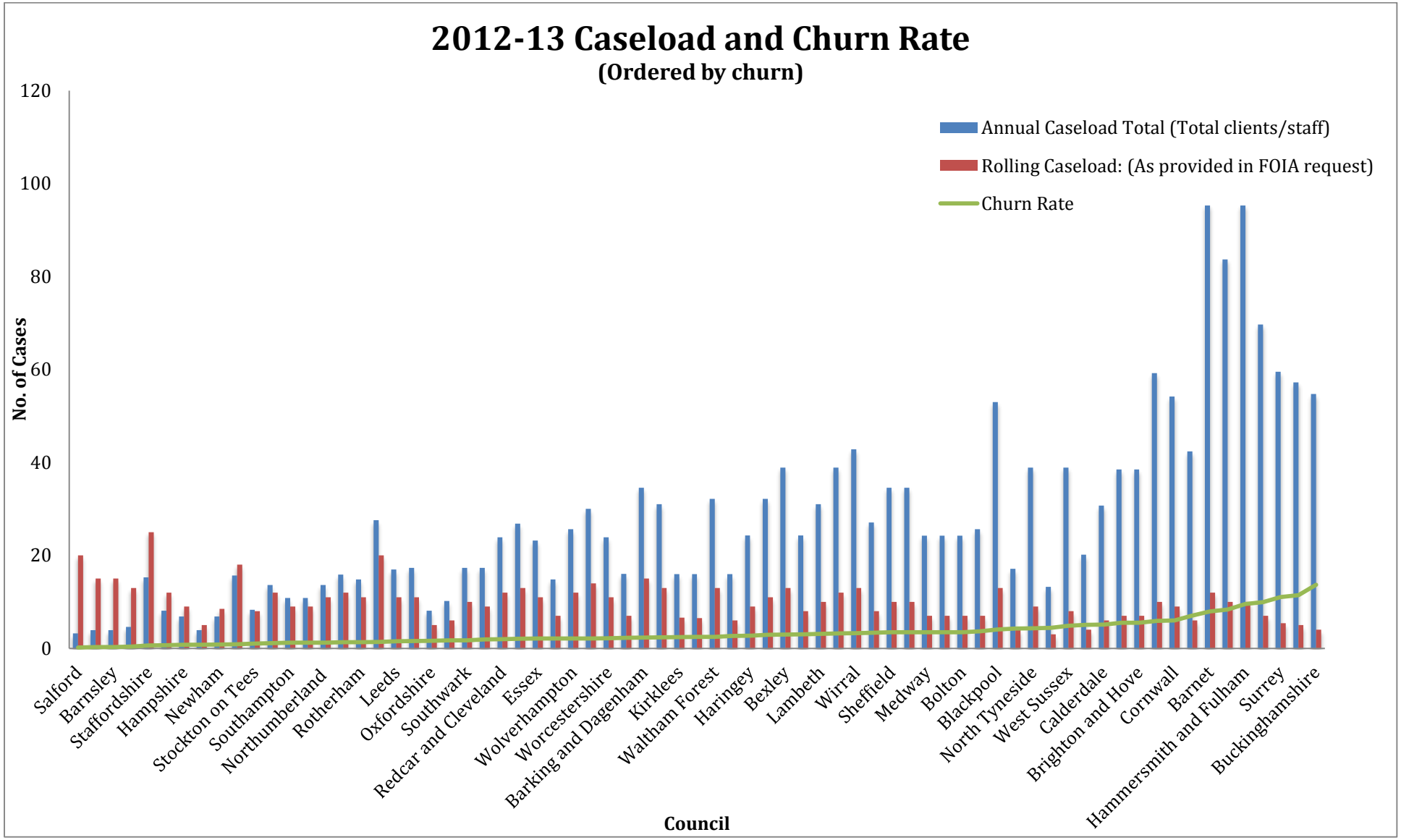
One of the clearest findings from the dataset was the consistent and significant discrepancy between caseload calculated according to staff and client numbers and caseload averages provided directly through FOIA responses.

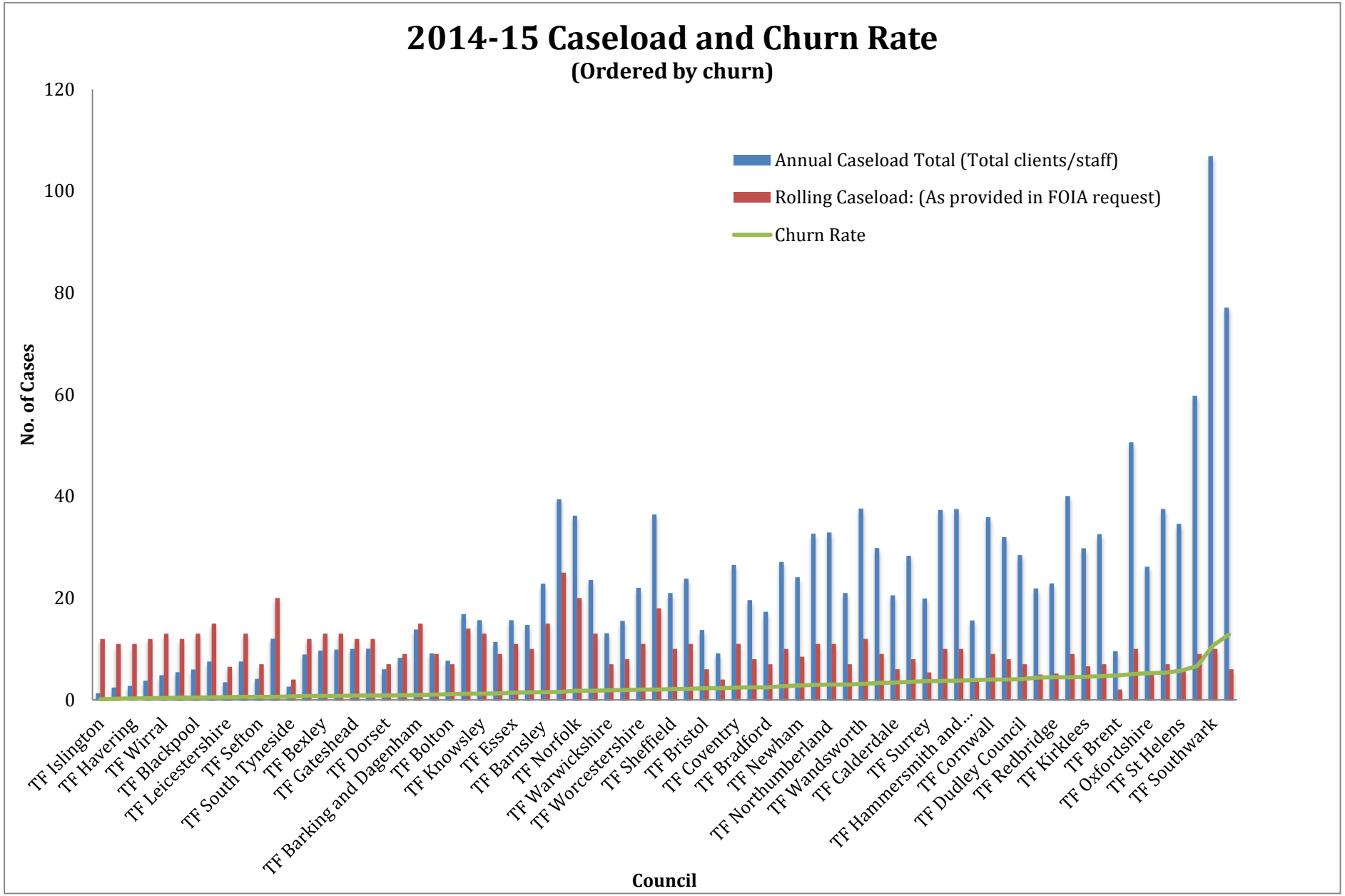
Caseload as calculated according to the figures released by the government in 2013-14 was 32 and in 2014-15 was 22. Average caseload cited in response to FOIA requests was the same across both years, and was a much lower 10.

There are different ways to account for this gap between the two caseload figures. The first is to presuppose that the higher caseloads calculated according to the aggregate figures in the government data releases represent an annual total caseload rather than the 'rolling' caseload held at any one time by a key worker. This might mean that a single worker on an average TF programme might have a caseload of 32 over the full course of the year, but would only be holding 10 or so active cases at any one time.

On this basis the dataset has a calculated 'churn' rate for the client turnover suggested by the discrepancies between these numbers. The average churn rate for 2013-14 is 4 and for 2014-15 is 3. This would mean that in 2013-14, the average key worker on a TF programme was working with 4 different rolling caseloads of clients, turning each caseload around within three months. The same average key worker would have slowed their pace a little in the following year, but would still have been turning around each rolling caseload within four months. At the upper end of the dataset we encounter churn rates that would translate to caseload turnaround every four or five weeks. This is a very fast pace.

Caseload levels and churn rates across both years are given in the charts below.

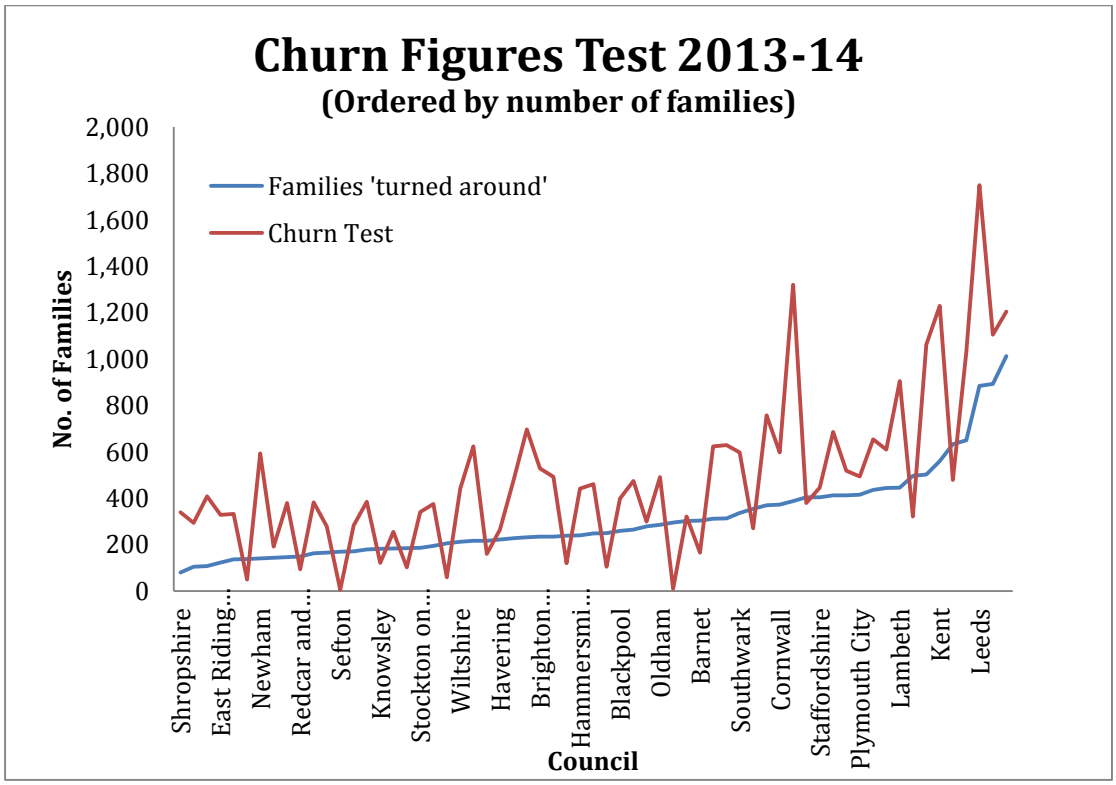

2012-13 Caseload and Churn Rate
(Ordered by churn)
Annual Caseload Total (Total clients/staff)
Rolling Caseload: (As provided in FOIA request)
Churn Rate
No. of Cases
120
100
80
60
40
20
0
Salford
Barnsley
Staffordshire
Hampshire
Newham
Stockton on Tees
Southampton
Northumberland
Rotherham
Leeds
Oxfordshire
Southwark
Redcar and Cleveland
Essex
Wolverhampton
Worcestershire
Barking and Dagenham
Kirklees
Waltham Forest
Haringey
Bexley
Lambeth
Wirral
Sheffield
Medway
Bolton
Blackpool
North Tyneside
West Sussex
Calderdale
Brighton and Hove
Cornwall
Barnet
Hammersmith and Fulham
Surrey
Buckinghamshire
Council


2014-15 Caseload and Churn Rate
(Ordered by churn)
120
100
80
60
40
20
0
No. of Cases
Annual Caseload Total (Total clients/staff)
Rolling Caseload: (As provided in FOIA request)
Churn Rate
TF Islington
TF Havering
TF Wirral
TF Blackpool
TF Leicestershire
TF Sefton
TF South Tyneside
TF Bexley
TF Gateshead
TF Dorset
TF Barking and Dagenham
TF Bolton
TF Knowsley
TF Essex
TF Barnsley
TF Norfolk
TF Warwickshire
TF Worcestershire
TF Sheffield
TF Bristol
TF Coventry
TF Bradford
TF Newham
TF Northumberland
TF Wandsworth
TF Calderdale
TF Surrey
TF Hammersmith and...
TF Cornwall
TF Dudley Council
TF Redbridge
TF Kirklees
TF Brent
TF Oxfordshire
TF St Helens
TF Southwark
Council

Calculating churn rates also allows us to cross-check the 'turned around' figures provided in the data releases. If the churn rates are an accurate reflection of the pace of work in TF programmes, then they ought to tally up with the number of families exiting the programme at the end of each year. So that, for instance, if the staff/client caseload is 32, and the rolling caseload is 10, 'turned around' figures should – on average – show that 2/3rds of families had exited the programme by the end of the year, reflecting the pace of turnover for the TF worker.

However when that calculation is carried out, we find that there is no clean correlation between churn rates and exit figures. The average number of families 'turned around' from the data releases for 2013-14 is 295, the number of families that ought to appear if the churn rates are a reflection of TF practice is 335.

The line graph below shows the mismatch between the two. We do see a rising line calculated using the churn rate, but it does not follow the published 'turned around' figures smoothly.



For 2014-15 the mismatch is even more striking. The average number of families 'turned around' according to the data release is 495, the average number according to the churn test is 109.

In essence in 2013-14, the first year of the programme, when staff levels are lower and caseloads higher, we have TF workers who are not – it would appear – churning their caseloads as efficiently as the numbers suggest that

they should. In 2014-15 something very odd seems to happen. TF programmes 'turn around' four times as many families – on average – as the average caseload figures suggest that they should.

It is hard to account for this but at the very least it points to some sort of problem with the aggregated figures provided in the data release. It could be, for instance, that the numbers of families 'turned around' are not very accurately divided between the two years of the programme, perhaps because of the difficulty of establishing an exit point clearly, and leading to a kind of 'double accounting' problem, or it could be that there are other inaccuracies at work. It is hard to see, however, how the data release figures and the FOIA figures could both be accurate.

4. Capacity

As well as analyzing caseload and estimating churn rates based on the figures, I used the data to calculate the 'capacity' of each TF programme. The capacity calculation is worked out on the basis that each key worker in a TF programme is equally distributing her hours of support across each one of the families in her caseload, essentially the average number of hours per person. Of course it is very unlikely that support would in reality be distributed in this way given that families will require differential levels of involvement and will have varied concerns that need to be addressed in highly tailored ways. The averaged capacity of a programme is therefore not taken here to be an accurate representation of the real distribution of support hours.

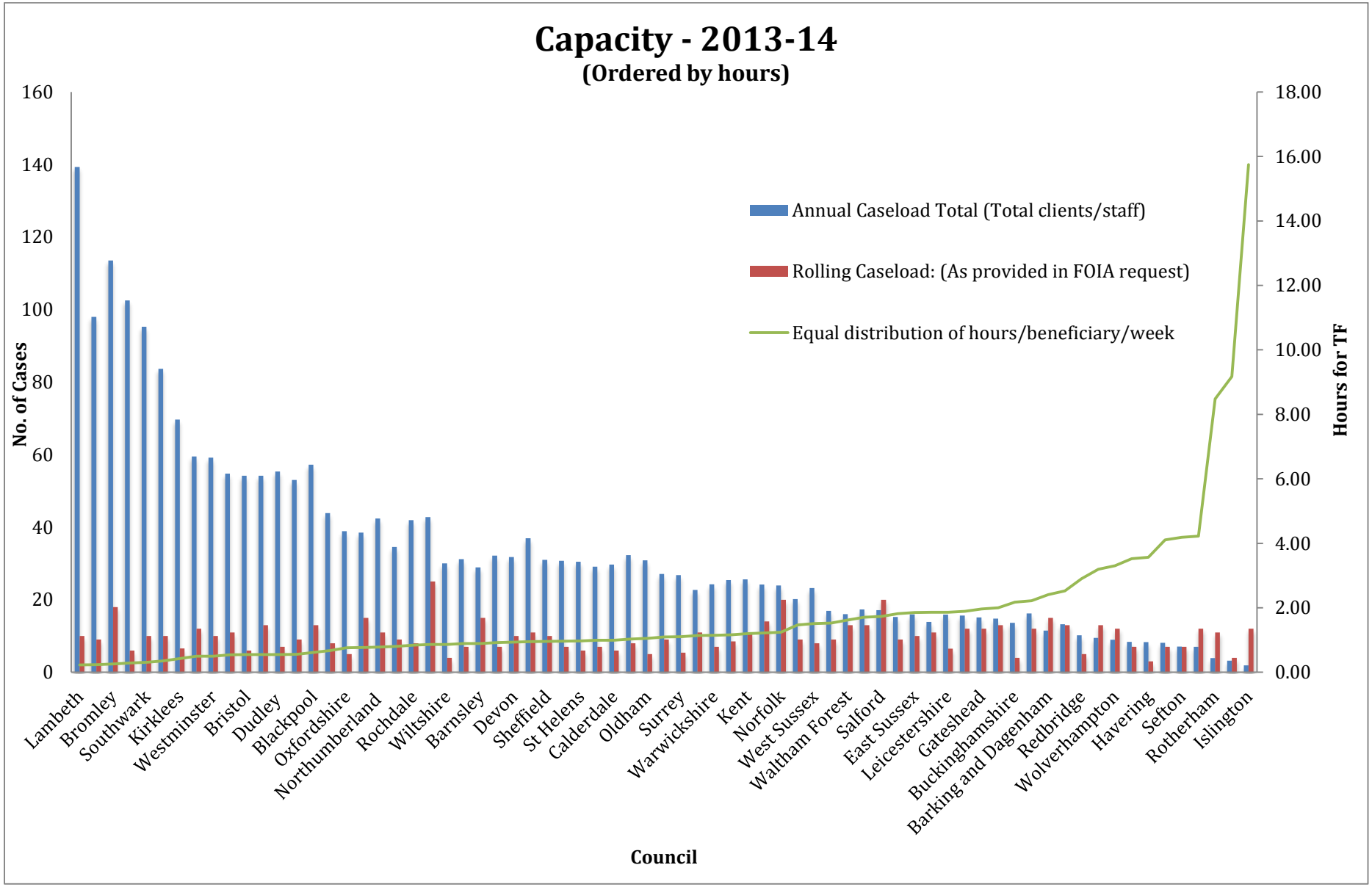
What it does give, however, is a sense of the maximum amount of time a key worker could give a family *if* each family *required* the same level of support. If one family requires more than this distributed amount of support, then other families will necessarily come in for less time unless another variable is changed, for instance the number of hours worked, or the time spent on admin tasks rather than direct contact. With that in mind, the dataset is set up to take into account the percentage of time a TF programme key worker typically spends on tasks that are not directly related to supporting families. This percentage estimate was taken from FOIA responses. Where no response was given to this question, a baseline of 20% was used based on Community Care's 2012 research with social workers.

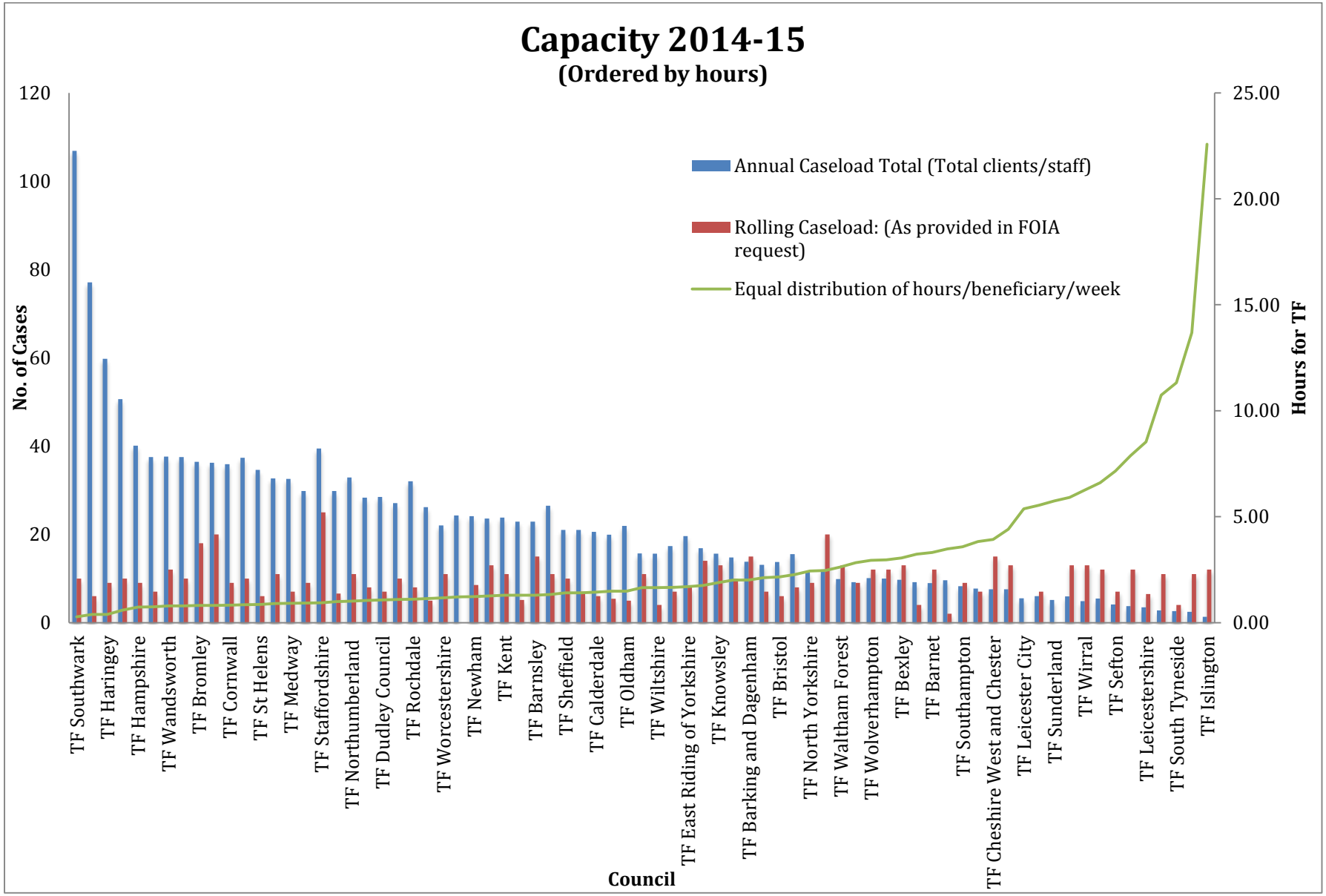
The dataset also assumes that all TF staff work the standard contracted 37 hours/week. It is quite likely that staff in fact work longer hours than these. This would affect the capacity calculation, perhaps significantly, but this is the kind of tweak that could easily be changed to reflect real worked hours in a further iteration of the model. It should be noted, however, that the FOIA requests specifically asked about overtime worked and where this question was answered there was no admission of extra hours or overtime paid to TF workers.

The hypothesis explored here is that the capacity calculation gets us a little deeper into the implications of caseload figures. Capacity could help to model the effects of changing aspects of a programme's delivery, such as admin and travel time for instance. It could be useful to see if this has a significant effect on the amount of time that a key worker is able to spend with a family. It is also a more nuanced way of assessing different programmes comparatively.

There were wide variations between programme capacity and the mean should be treated with some caution as a result, but it is nevertheless interesting. In 2013-14 the mean number of hours that a TF key worker could spend on each client family's case – assuming an equal distribution as outlined above – was 2 hours a week. In 2014-15 this had gone up to 3 hours a week. This average capacity was typically correlated with a caseload of around 15 families. A caseload of 44, by contrast, reduced contact time to around 1 hour a week.

The charts below map caseload against weekly capacity across both years.


Capacity - 2013-14
(Ordered by hours)
No. of Cases
Hours for TF
Council
Annual Caseload Total (Total clients/staff)
Rolling Caseload: (As provided in FOIA request)
Equal distribution of hours/beneficiary/week
0
20
40
60
80
100
120
140
160
0.00
2.00
4.00
6.00
8.00
10.00
12.00
14.00
16.00
18.00
Lambeth
Bromley
Southwark
Kirklees
Westminster
Bristol
Dudley
Blackpool
Oxfordshire
Northumberland
Rochdale
Wiltshire
Barnsley
Devon
Sheffield
St Helens
Calderdale
Oldham
Surrey
Warwickshire
Kent
Norfolk
West Sussex
Waltham Forest
Salford
East Sussex
Leicestershire
Gateshead
Buckinghamshire
Barking and Dagenham
Redbridge
Wolverhampton
Havering
Sefton
Rotherham
Islington


Capacity 2014-15
(Ordered by hours)
No. of Cases
Hours for TF
Council
Annual Caseload Total (Total clients/staff)
Rolling Caseload: (As provided in FOIA request)
Equal distribution of hours/beneficiary/week
120
100
80
60
40
20
0
25.00
20.00
15.00
10.00
5.00
0.00
TF Southwark
TF Haringey
TF Hampshire
TF Wandsworth
TF Bromley
TF Cornwall
TF St Helens
TF Medway
TF Staffordshire
TF Northumberland
TF Dudley Council
TF Rochdale
TF Worcestershire
TF Newham
TF Kent
TF Barnsley
TF Sheffield
TF Calderdale
TF Oldham
TF Wiltshire
TF East Riding of Yorkshire
TF Knowsley
TF Barking and Dagenham
TF Bristol
TF North Yorkshire
TF Waltham Forest
TF Wolverhampton
TF Bexley
TF Barnet
TF Southampton
TF Cheshire West and Chester
TF Leicester City
TF Sunderland
TF Wirral
TF Sefton
TF Leicestershire
TF South Tyneside
TF Islington

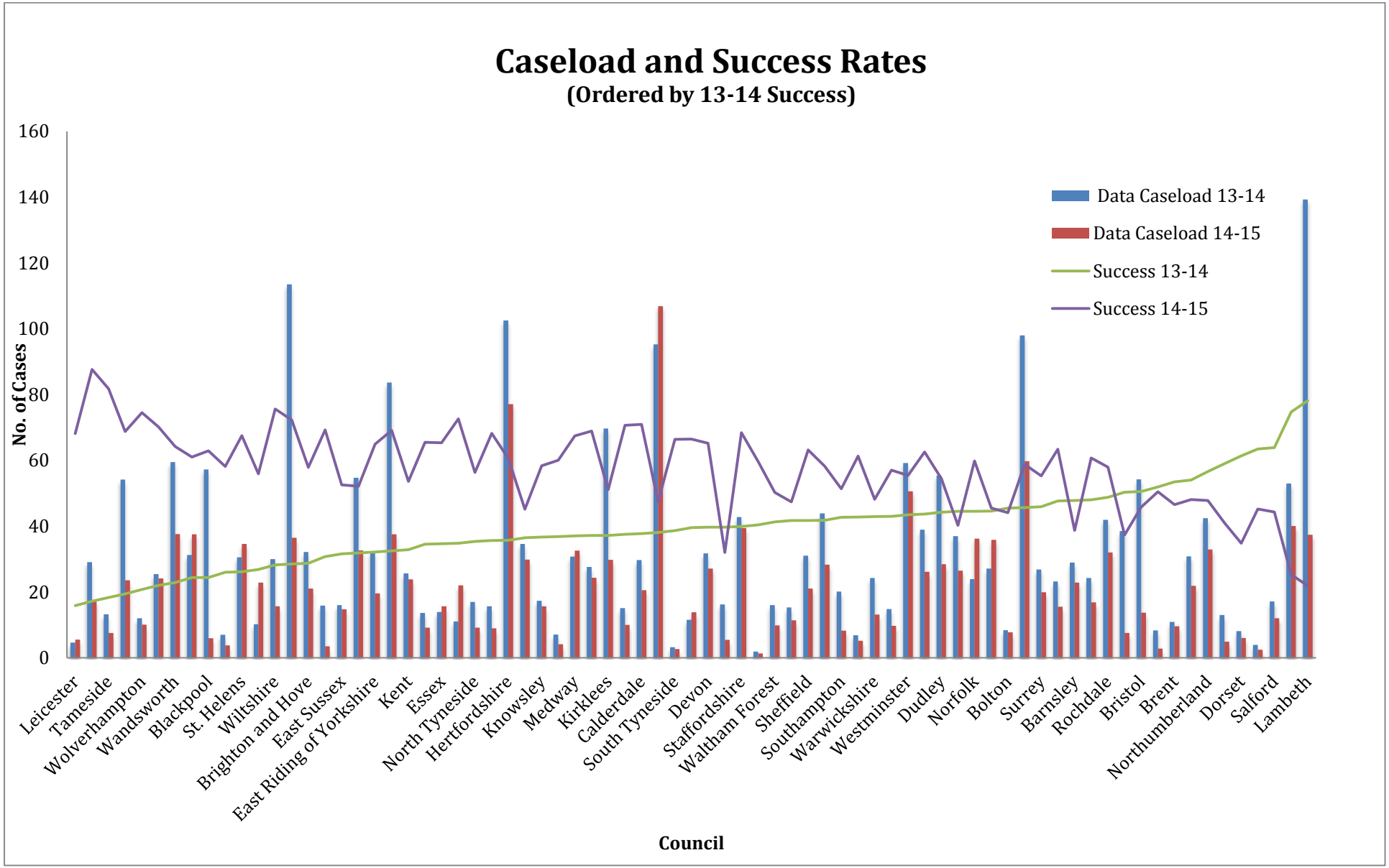
5. Success

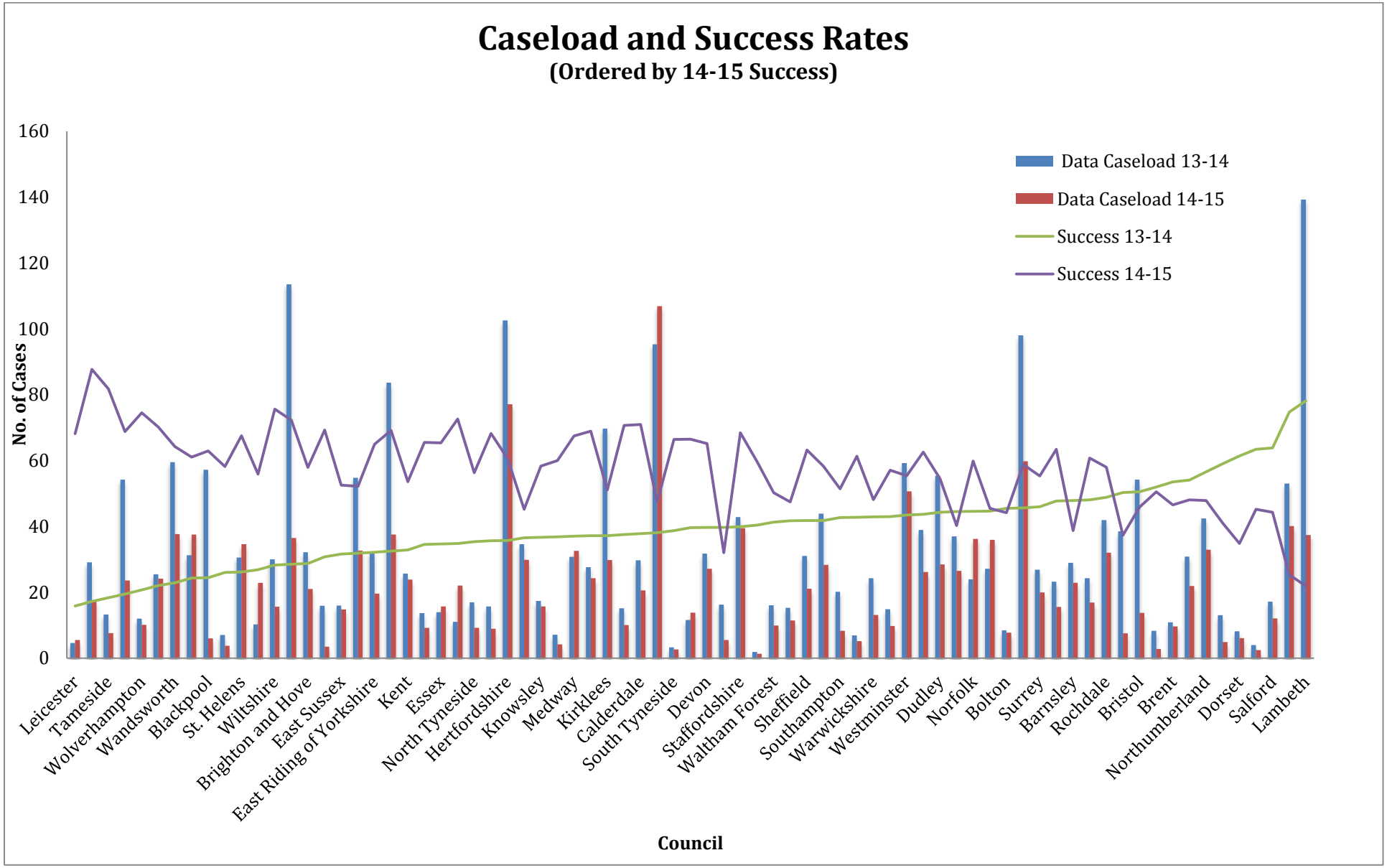
The final calculation in the TF dataset is the 'success' rating. This was based on the turned around figures collected through both the government data releases (and FOIA in the second round of queries) as a percentage of the total number of families 'worked with' during the year. The intention here was again to test the 'turned around' numbers and see whether there was any discernible difference in 'success' between the two years for which data was available, and to look for any correlations between 'success' ratings and other aspects of programme delivery, notably caseload and programme capacity.

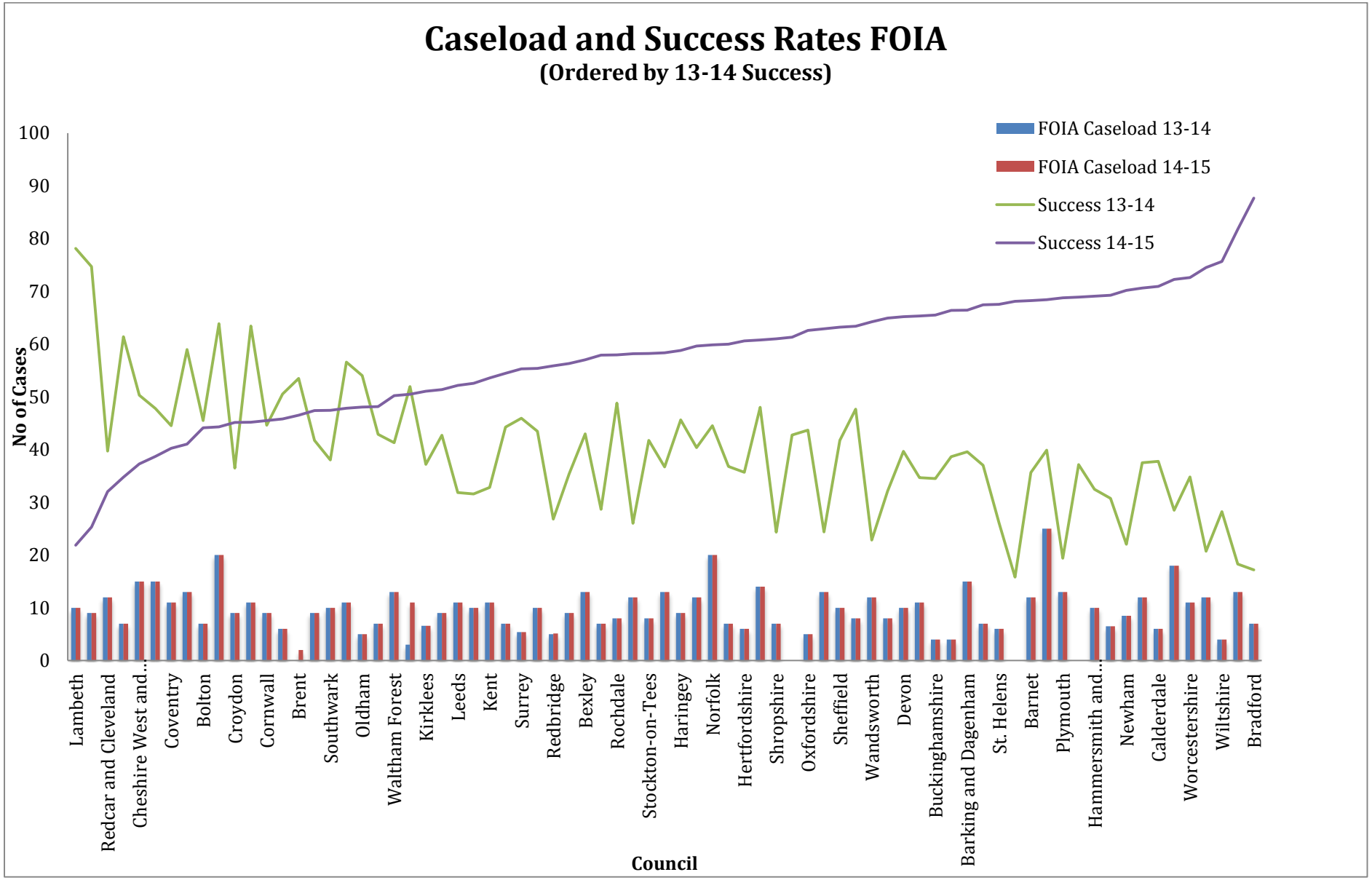
The first finding is that as caseloads come down and capacity goes up overall in 2014-15, we do find, as we might expect, a boost in 'success' rates. These gave an average 'success' rating of 40% for 2013-14 and 57% for 2014-15.

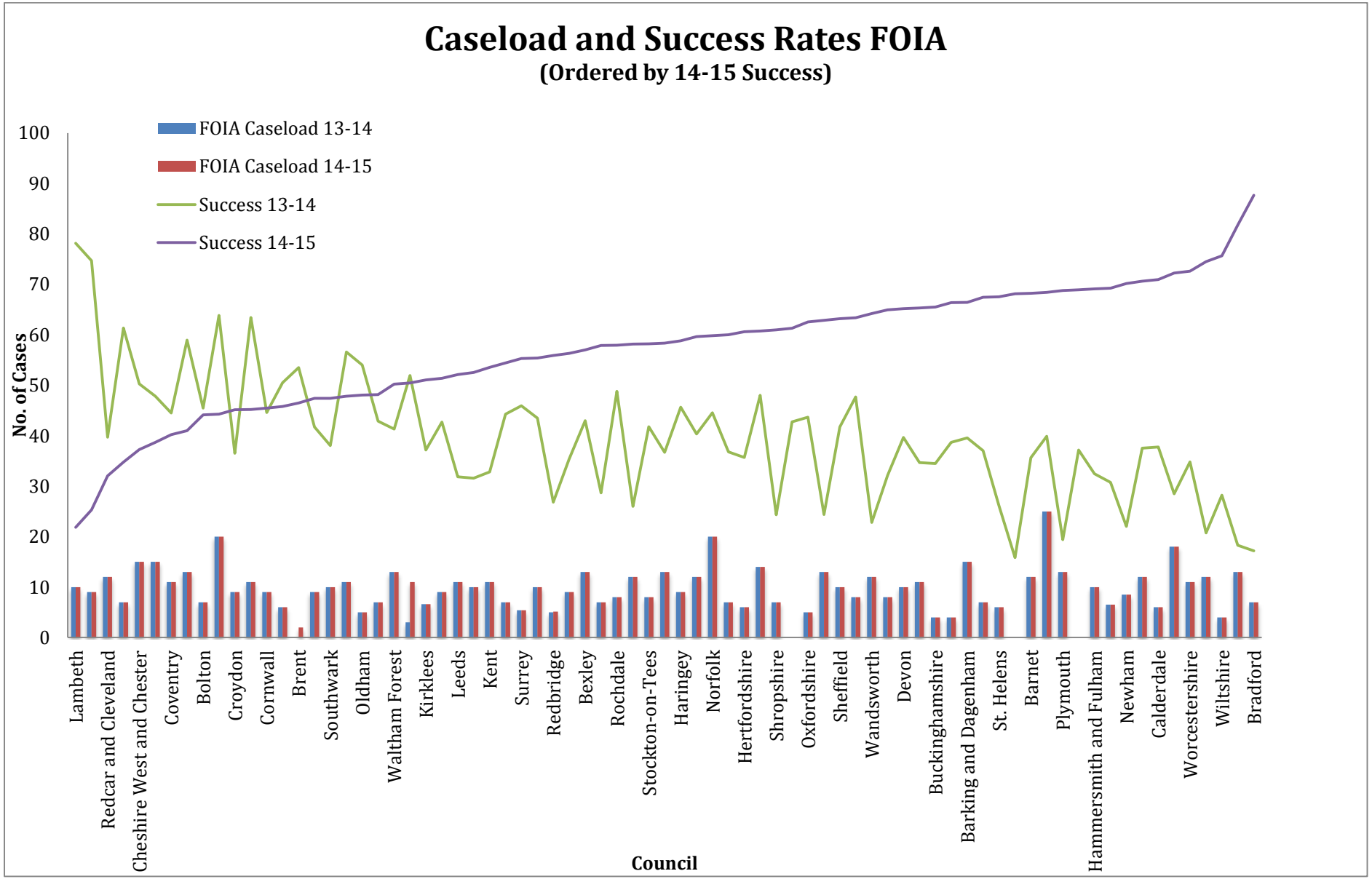
When it comes to correlations of caseload, capacity and success more generally however, there is no very clearly discernible pattern (as can be seen in the charts below). The churn rate test results as outlined above already show that 'turned around' numbers, and consequently success ratings, should be treated with caution. It is not therefore surprising to find that caseload and success do not seem to be correlated in any straightforward way. This absence of correlation persists even when the caseload figures shared in the FOIA responses are used instead of the staff/client calculated figure.

It is probable that the 'turned around' figures in the government releases are the least reliable of all the data points shared.


Caseload and Success Rates
(Ordered by 13-14 Success)
No. of Cases
160
140
120
100
80
60
40
20
0
Data Caseload 13-14
Data Caseload 14-15
Success 13-14
Success 14-15
Leicester
Tameside
Wolverhampton
Wandsworth
Blackpool
St. Helens
Wiltshire
Brighton and Hove
East Sussex
East Riding of Yorkshire
Kent
Essex
North Tyneside
Hertfordshire
Knowsley
Medway
Kirklees
Calderdale
South Tyneside
Devon
Staffordshire
Waltham Forest
Sheffield
Southampton
Warwickshire
Westminster
Dudley
Norfolk
Bolton
Surrey
Barnsley
Rochdale
Bristol
Brent
Northumberland
Dorset
Salford
Lambeth
Council


Caseload and Success Rates
(Ordered by 14-15 Success)
No. of Cases
160
140
120
100
80
60
40
20
0
Data Caseload 13-14
Data Caseload 14-15
Success 13-14
Success 14-15
Leicester
Tameside
Wolverhampton
Wandsworth
Blackpool
St. Helens
Wiltshire
Brighton and Hove
East Sussex
East Riding of Yorkshire
Kent
Essex
North Tyneside
Hertfordshire
Knowsley
Medway
Kirklees
Calderdale
South Tyneside
Devon
Staffordshire
Waltham Forest
Sheffield
Southampton
Warwickshire
Westminster
Dudley
Norfolk
Bolton
Surrey
Barnsley
Rochdale
Bristol
Brent
Northumberland
Dorset
Salford
Lambeth
Council


Caseload and Success Rates FOIA
(Ordered by 13-14 Success)
FOIA Caseload 13-14
FOIA Caseload 14-15
Success 13-14
Success 14-15
No of Cases
100
90
80
70
60
50
40
30
20
10
0
Lambeth
Redcar and Cleveland
Cheshire West and...
Coventry
Bolton
Croydon
Cornwall
Brent
Southwark
Oldham
Waltham Forest
Kirklees
Leeds
Kent
Surrey
Redbridge
Bexley
Rochdale
Stockton-on-Tees
Haringey
Norfolk
Hertfordshire
Shropshire
Oxfordshire
Sheffield
Wandsworth
Devon
Buckinghamshire
Barking and Dagenham
St. Helens
Barnet
Plymouth
Hammersmith and...
Newham
Calderdale
Worcestershire
Wiltshire
Bradford
Council


Caseload and Success Rates FOIA
(Ordered by 14-15 Success)
FOIA Caseload 13-14
FOIA Caseload 14-15
Success 13-14
Success 14-15
No. of Cases
100
90
80
70
60
50
40
30
20
10
0
Lambeth
Redcar and Cleveland
Cheshire West and Chester
Coventry
Bolton
Croydon
Cornwall
Brent
Southwark
Oldham
Waltham Forest
Kirklees
Leeds
Kent
Surrey
Redbridge
Bexley
Rochdale
Stockton-on-Tees
Haringey
Norfolk
Hertfordshire
Shropshire
Oxfordshire
Sheffield
Wandsworth
Devon
Buckinghamshire
Barking and Dagenham
St. Helens
Barnet
Plymouth
Hammersmith and Fulham
Newham
Calderdale
Worcestershire
Wiltshire
Bradford
Council

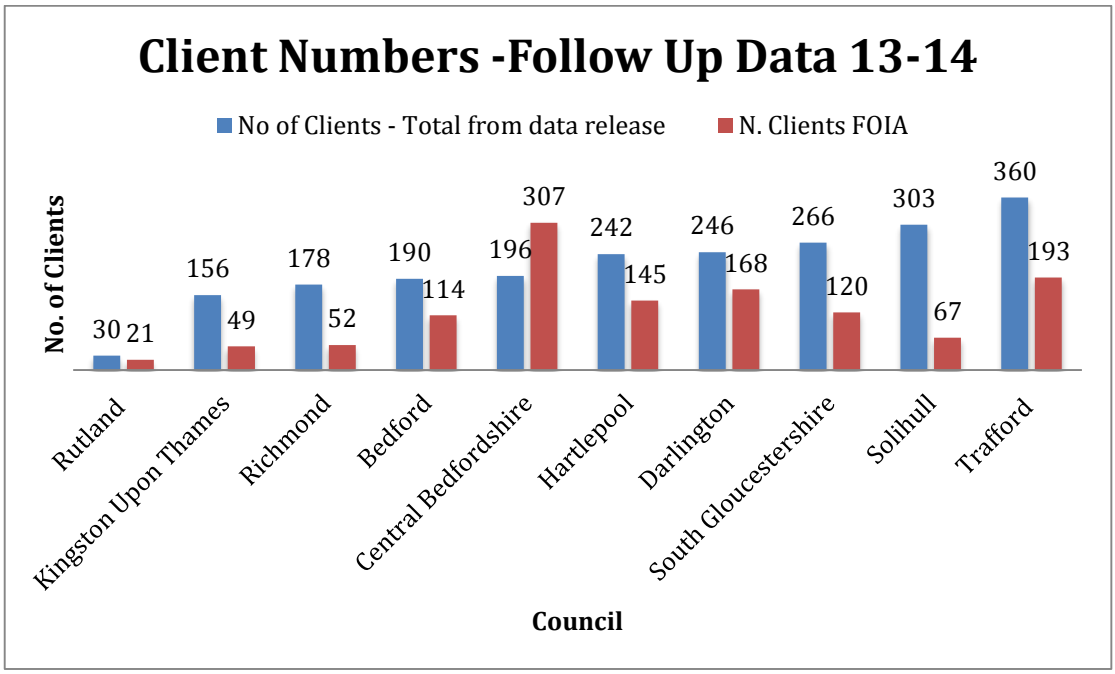
6. Exploring Discrepancies

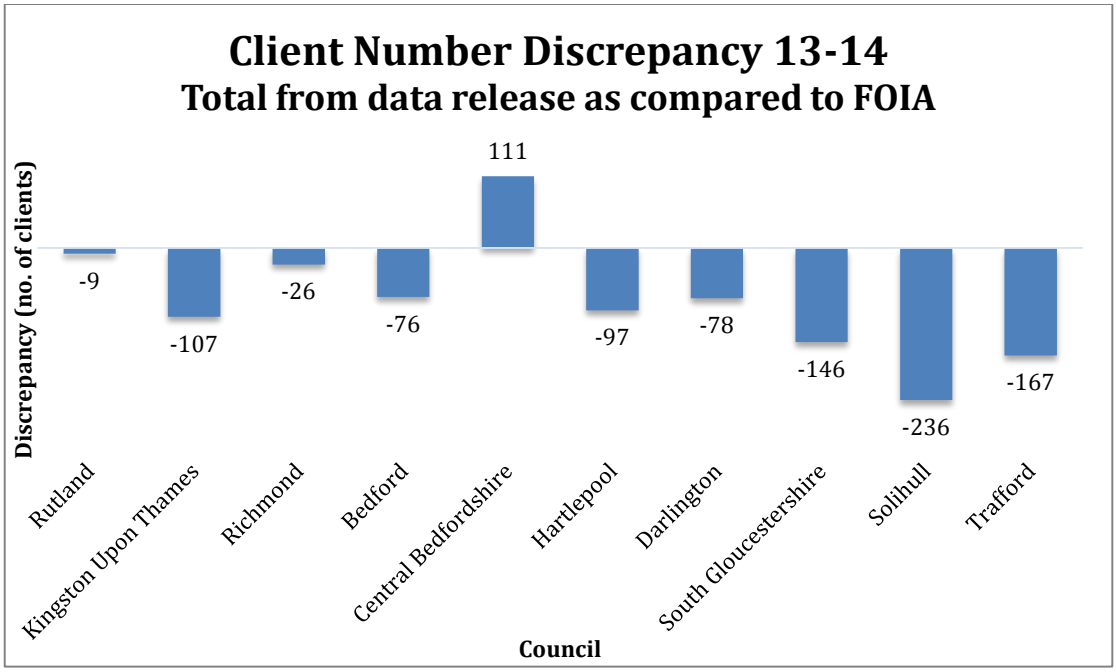
As noted above there were very significant variations between the numbers published by LAs in the government data releases and the equivalent numbers released in response to my FOIA requests.

In the second round of FOIA queries I sought to explore discrepancies by asking more specific questions about the numbers of families 'worked with' and 'turned around'.

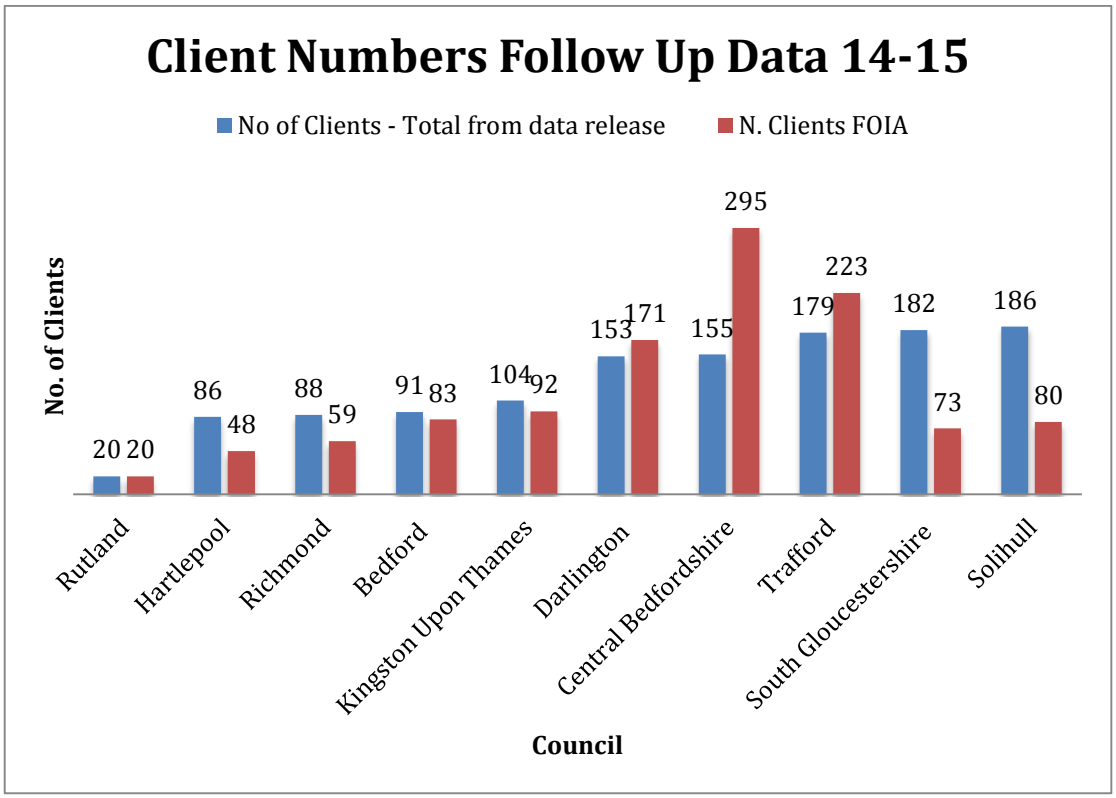
Perhaps unsurprisingly, numbers of families 'worked with' differed between data releases and FOIA responses. For the 13-14 data there was not a single Local Authority that provided usable data that gave the same figure for clients in the FOIA response as had been given through the government data release. The average number of families from the data release was 206, but this figure was a much lower 124 as calculated from FOIA responses.

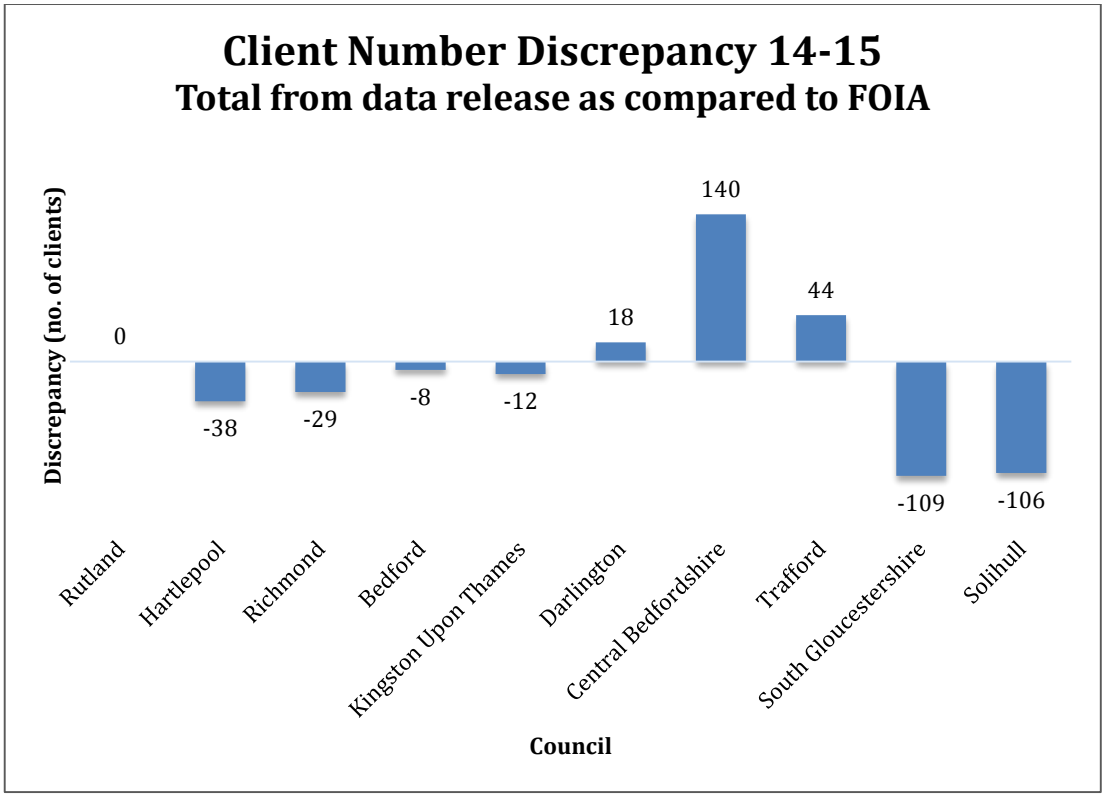
The bar chart below shows the differences between data release and FOIA client numbers for 13-14. FOIA figures were lower for each Local Authority except in the case of Central Bedfordshire:



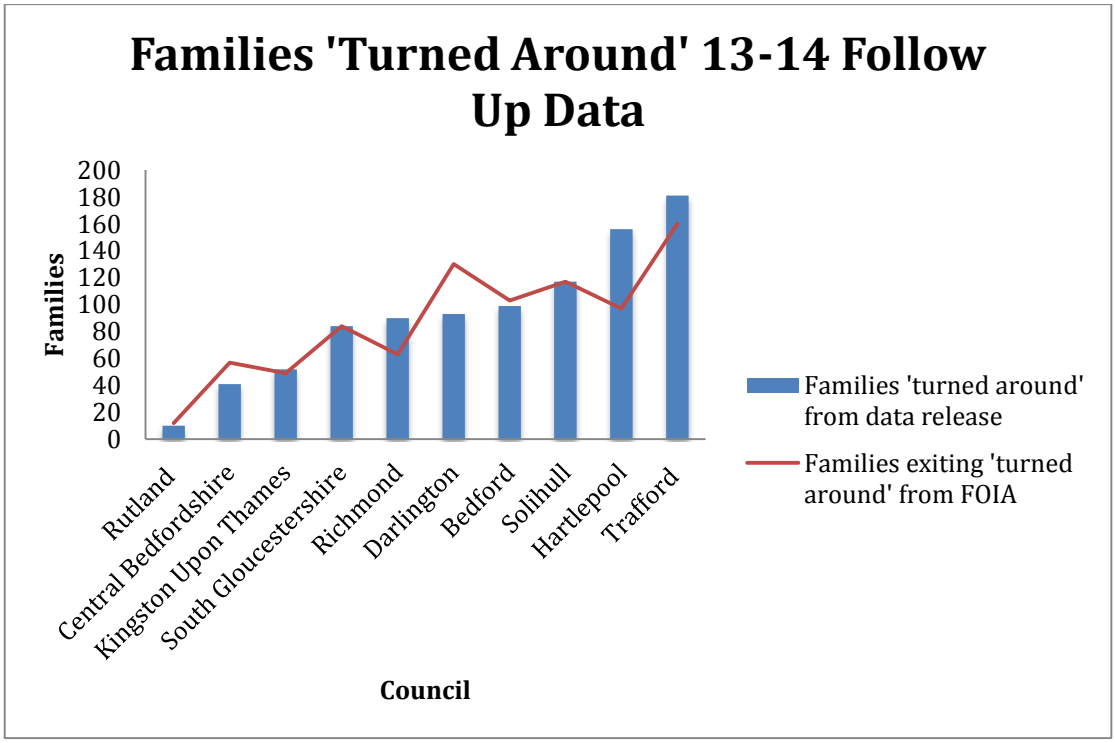


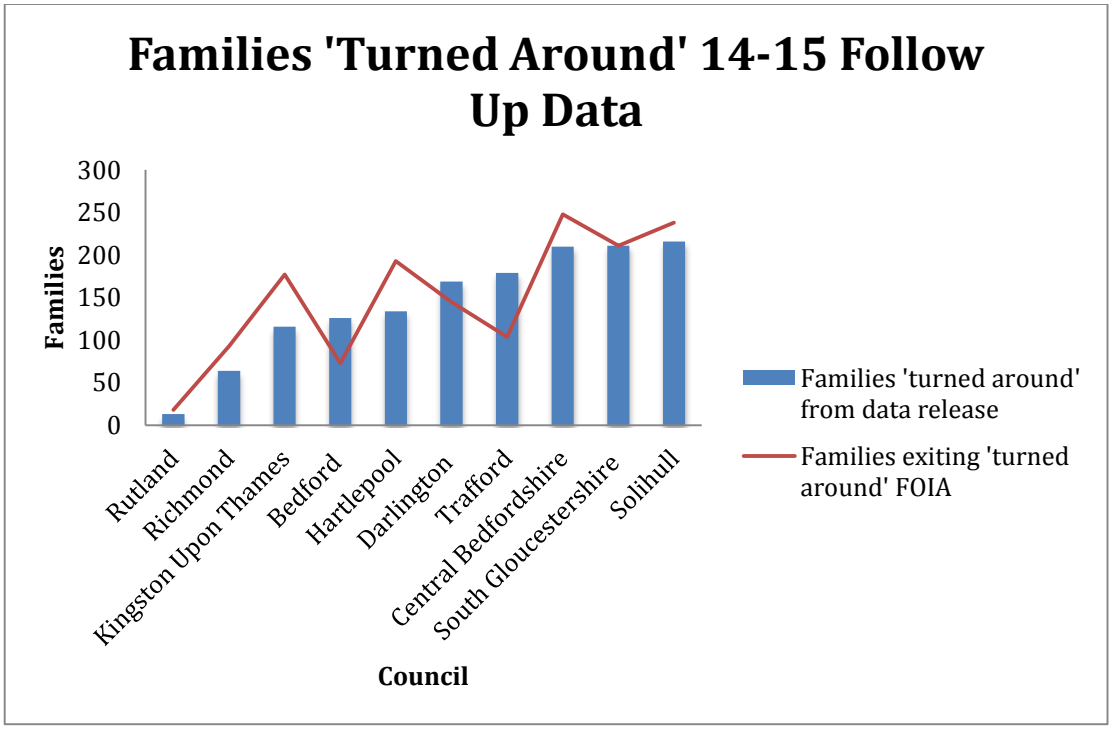
There were discrepancies in the 14-15 data too, as the bar chart below shows, but in this case FOIA numbers exceeded data release numbers for three Local Authorities, with Central Bedfordshire once again sharing much higher client numbers through their FOIA response.





Similar discrepancies existed in the numbers of families 'turned around', although these were less significant in both years. The average number of families 'turned around' in the data releases for 13-14 in the second dataset was 92; in the FOIA responses it was 87. For 14-15, the average was 144 based on the data releases, and 150 based on FOIA. Bar charts for both years show the values for each Local Authority:





The second dataset was also used to explore caseload, capacity and success ratings taking discrepancies into account.

In this second much smaller dataset, annual client/staff caseload was much lower across both years, with an average of 20 in 13-14 and of 11 in 14-15, again showing the reduction in caseload in the second year of the programme. In line with the caseload figures, capacity was higher than the dataset average from the first round of queries in the follow up data, with a maximal distribution of 3 hours/week in 13-14 and close to 5 hours/week in 14-15.

Given the much lower caseload average and higher programme capacity, it was interesting to see whether success rates would show any significant variation from the initial larger dataset. For 13-14, the data matched very closely with the average from the first round, with a 41% success rate. The second year, however, showed very surprising results, with families 'turned around' exceeding the client total from the data release in 7 of the 10 new data points. This, clearly, could not accurately represent the programmes in question given that this would imply that TF workers were turning around more families than were registered with them, an odd result that we already encountered in the churn test in the larger datasets.

As already noted, this finding may be due to the aggregation of figures across the two years for which the government has shared data on the TF programmes. This may have led to some 'double accounting' where figures 'turned around' in the 2013-14 data releases, reappeared in the 2014-15 numbers as part of an aggregated whole.

As part of the analysis of the data, I subtracted 'turned around' figures from 13-14 from the 14-15 releases, but this may not have been a sufficient measure to ensure that the 'true' figures were used to calculate caseload, capacity and success rates.

Even after careful analysis, the discrepancies subsist and cannot readily be explained away.

Conclusions

1. Despite very clear problems with the data retrieved through the FOIA requests and the government data releases, some interesting findings came from this analysis. The data strongly suggests that:

    a. The political 'colour' of the local council has an effect on TF staffing levels
    b. TF programmes have high sickness rates, above the national average in most cases
    c. Staffing levels and caseload are not directly correlated with sickness rates
    d. There is, however, a correlation between sickness rates and council political leadership

2. More importantly, the TF dataset suggests that there are serious weaknesses in the data released by the government concerning delivery of the TF programmes.

    Whilst it is possible that there is no straightforward correlation between caseload, programme capacity and the 'turning around' of families through the TF approach, this seems unlikely. Added to the consistent discrepancies between data release and FOIA numbers it seems more probable that data release figures are inaccurate, misleading, distorted or inflated.

    Inflation seems particularly likely given the very different figures supplied on average caseload through the FOIA responses. Given the bizarre results from the churn tests, it is hard to believe that the FOIA caseload averages – which are directly in line with DCLG guidance on the TF programme – should be less accurate than the calculated caseload that comes from the aggregated figures.

    It may be that this inflation is a simple result of the aggregation of figures across the two years of the programme. It may be –as in the case of the FOIA figures – that differential definitions of outcomes, and variable monitoring have produced surprising, or inaccurate, results, and particularly that staffing levels were regularly underestimated.

    All the same it seems at least reasonable to conclude that these tables do not give us a reliable and transparent account of the working of the programmes across England and Wales.

    They are not as good as they could be.

There is a great deal wrong with the TF dataset as it is presented here, but it may all the same point to ways of carrying out comparable assessment of social programmes that would not be reliant on aggregated figures, and, that could – for

instance – usefully model innovations in existing programmes in advance of a pilot, inform those seeking to begin a new programme, and offer comparative information that would be valuable to all.

This is not enough on its own, but it does provide a starting point for a more detailed approach to data collection, publication and analysis. It is something on which to build.

Annex

The following councils refused, on the grounds described, to provide data:

Information Not Held	Too Expensive
Camden Borough Council	Ealing Borough Council
Cheshire East Council	Hackney Borough Council
Derbyshire County Council	Lancashire County Council
Hillingdon Borough Council	Merton Borough Council
Cambridgeshire County Council	Thurrock Borough Council
Lewisham Borough Council	Tower Hamlets Borough Council
Lincolnshire County Council	Walsall Metropolitan Borough Council
Luton Borough Council	
Manchester City Council	
North East Lincolnshire Council	
Nottinghamshire County Council	
Reading Borough Council	
Slough Borough Council	
Stockport Metropolitan Borough Council	
Suffolk County Council	
Wakefield City Council	